On the Capacity of Certain Space-Time Coding Schemes

Constantinos B. Papadias

Global Wireless Systems Research Department, Bell Laboratories, Lucent Technologies, 791 Holmdel-Keyport Road, Holmdel, NJ 07733, USA
Email: papadias@lucent.com

Gerard J. Foschini

Wireless Communications Research Department, Bell Laboratories, Lucent Technologies, 791 Holmdel-Keyport Road, Holmdel, NJ 07733, USA
Email: gjf@lucent.com

We take a capacity view of a number of different space-time coding (STC) schemes. While the Shannon capacity of multiple-input multiple-output (MIMO) channels has been known for a number of years now, the attainment of these capacities remains a challenging issue in many cases. The introduction of space-time coding schemes in the last 2–3 years has, however, begun paving the way towards the attainment of the promised capacities. In this work we attempt to describe what are the attainable information rates of certain STC schemes, by quantifying their inherent capacity penalties. The obtained results, which are validated for a number of typical cases, cast some interesting light on the merits and tradeoffs of different techniques. Further, they point to future work needed in bridging the gap between the theoretically expected capacities and the performance of practical systems.

Keywords and phrases: MIMO systems, space-time coding, Bell labs Layered Space Time (BLAST), space-time spreading (STS), channel capacity, space-time processing (STP), transmit diversity.

1. INTRODUCTION

The combined use of antenna arrays and sophisticated multiple-input multiple-output (MIMO) transceiver techniques has boosted the anticipated spectral efficiencies of wireless links in the last five years or so. The MIMO channel capacity expressions derived in [1] indicate that the spectral efficiencies of MIMO channels can grow approximately linearly with the (minimum of the) number of antennas available on each side of the link.

Similar to the case of single-input single-output (SISO) channels, the attainment of the theoretically promised capacities in practice, has to rely on strong encoding/decoding techniques. In the SISO case, it took about fifty years to approach closely (with the advent of Turbo codes [2]) the channel capacities predicted by Shannon [3]. In the MIMO case, an initial (the so-called D-BLAST) architectural superstructure was proposed in [1] that is theoretically capable of achieving the channel capacity. The quest for *practical* capacity-approaching STC techniques is however ongoing. Interestingly, it seems that some existing STC techniques [4, 5], already allow to approach closely the channel capacities in a number of cases [6, 7]. These quite rapid advancements are of course not unrelated to the mature state of SISO encoding/decoding techniques. At the same time, there still

exist many cases of interest, where more research is needed in order to approach the capacities of MIMO systems in practice.

In this paper, we will attempt to quantify the performance of certain STC techniques, in terms of their "attainable" capacities. By "attainable capacities" we mean the capacities achieved by different techniques with the use of progressively stronger *known* (typically SISO) encoding/decoding techniques. In other words, we will quantify the *irreducible capacity penalties* inherent in certain STCs, due to the way they process signals at the transmitter, as well as at the receiver. Our general framework targets STC techniques which, when viewed end-to-end, can be broken down to a number of SISO problems. This will allow the evaluation of attainable capacities by quantifying the spectral efficiency of each component SISO channel. We then present a number of STC techniques that fit well within our framework and derive their capacities. The evaluation of these capacities helps not only compare some of the existing techniques, but also to identify cases where further research is needed.

The remainder of the paper is organized as follows. In Section 2, we provide our working assumptions, as well as some background on the topic. In Section 3, we define the notion of decomposable STCs, as well as some other relevant features of such codes. In Section 4, we focus on a num-
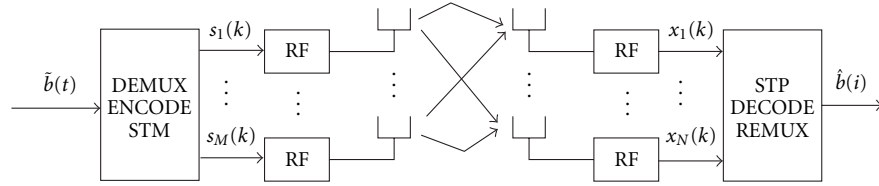
FIGURE 1: A generic (M, N) multiple antenna system.

ber of recently developed STC techniques for multiple-input single-output (MISO) channels, and show what are their attainable capacities. In Section 5, we present similar results for MIMO systems. In Section 6, we show some numerical results for outage capacities of MIMO Rayleigh-faded channels in a number of cases of interest. Finally, in Section 7 we present our conclusions, as well as some directions for future work.

2. BACKGROUND AND ASSUMPTIONS

Figure 1 shows a generic architecture of a wireless system with M transmitter and N receiver antennas. Such a system will be denoted in the remainder of the paper as (M, N). The continuous-time input stream $\tilde{b}(t)$ is assumed to be carrying the original primitive data stream $\{\tilde{b}(i)\}$ that is to be communicated to the receiver. The input stream is then processed by the shown DEMUX/ENCODE/STM unit, whose output is an ensemble of M parallel data streams, each one of which is separately upconverted and transmitted over the MIMO channel.

The DEMUX/ENCODE/STM unit includes the following operations:

(1) demultiplexing;
(2) encoding;
(3) spatial multiplexing.

These operations may be ordered differently and can be done in a more or less joint fashion. For example, the original bit stream may be encoded first as a whole, and then demultiplexed onto the M antennas. Alternatively, $\{\tilde{b}(i)\}$ may be first demultiplexed onto a number of sub-streams, each one of which is afterwards separately encoded independently. Either way, the encoded/demultiplexed sub-streams are then mapped through the so-called spatial multiplexer onto the M antennas for transmission. This mapping may be a simple 1-1 streaming of each encoded sub-stream on each antenna (such as in the original so-called V-BLAST transmission mentioned in [8]), or a more complex spatial mapping. At the receiver, after the signals are received with an antenna array, they are first converted to baseband. Then, they are processed in space and time (STP), decoded, and remultiplexed in the STP/DECODE/REMUX unit (again the order of these operations may be arbitrary). These operations attempt to recover as reliably as possible a replica of the original primitive bit stream $\{\tilde{b}(i)\}$. For the purposes of this paper, temporal interleaving is not explicitly accounted for, but it can be easily accomodated.

The way in which we will view MIMO systems throughout the paper is the following. The specific way in which the operations at the transmitter and the receiver mentioned above take place, imposes a number of constraints to the problem of achieving the MIMO capacity. We refer to the system that results after the imposition of these constraints as an architectural "STC super-structure." Each STC superstructure then admits a whole class of specific STCs, by applying different types of temporal error-correction codes. Our goal will be to identify what are the capacity penalties inherent to these STC super-structures, due to the imposition of constraints on both the transmitter and the receiver. Said differently, we will attempt to quantify the Shannon capacities that are attainable in each case.

Now we define the notation and assumptions that will be used throughout the rest of the paper. After error correction coding, interleaving, and demultiplexing (irrespective of the order into which these operations occur), the original bit stream $\{\tilde{b}(i)\}$ is converted to a number, say Q, of encoded sub-streams, denoted as $\{b_1(k)\}, \ldots, \{b_Q(k)\}$. Note that the number of encoded sub-streams Q will most often equal the number of transmitter antennas M, however this may not always be the case. Finally, the Q sub-streams are mapped through a spatial multiplexing operation, as shown in Figure 1, to the M sub-streams that are transmitted from the M antennas. We denote the sub-stream transmitted from the mth antenna by $\{s_m(k)\}$. We assume that the physical channel between the mth transmitter and the nth receiver antenna is flat-faded in frequency, it can be hence represented, at baseband, through the complex scalar h_{nm}. The baseband received signal at the receiver antenna array is then represented by the following familiar (narrow-band) mixing model:

$$\mathbf{x}(k) = \mathbf{H}\mathbf{s}(k) + \mathbf{n}(k), \tag{1}$$

where the involved quantities are defined as follows:

- $\mathbf{s}(k) := [s_1(k) \ \cdots \ s_M(k)]^T$ is the $M \times 1$ vector snapshot of transmitted sub-streams, each assumed of equal variance σ_s^2;
- $\mathbf{H}$ is the $N \times M$ channel matrix;
- $\mathbf{x}(k)$ is the $N \times 1$ vector of received signal snapshots;
- $\mathbf{n}(k)$ is the $N \times 1$ vector of additive noise samples, assumed i.i.d. and mutually independent, each of variance σ_n^2.

We also denote by the superscripts $*$, T, $\dagger$ the complex conjugate, transpose, and Hermitian transpose, respectively, of a scalar or matrix. The open-loop Shannon capacity of the

(M, N) flat-faded channel is given (see [1]) by the now familiar (so-called "log-det") formula:

$$C = \log_2 \left\{ \det \left(I_N + \frac{\rho}{M} HH^\dagger \right) \right\} \quad \text{[bps/Hz]}, \qquad (2)$$

where $\rho = M\sigma_s^2/\sigma_n^2$. This formula assumes the transmitter is constrained to communicate using i.i.d. random processes of equal power from each of the M antennas. Later we will refine the context to fully accommodate an open loop channel outage mode. As mentioned above, the capacity in (2) can be only achieved with the use of strong encoding (STC) techniques. In the remainder of the paper, we will attempt to quantify how much of this capacity is allowed to be attained within certain STC super-structures.

3. FEATURES OF STC ARCHITECTURAL SUPER-STRUCTURES

When viewing (1), the only visible imposed constraints are the equality between the powers of each sub-stream, the independent equal-power noise, and the flat channel characteristic. The absence of additional constraints would allow, in theory, the attainable capacity of the (M, N) system to be given by (2). The imposition of further constraints though, reflecting operations at both the transmitter and the receiver, may reduce the capacity in (2). We call this (potentially) reduced capacity the *constrained capacity* of a given architectural super-structure.

At the receiver, the received encoded vector signal $\mathbf{x}(k)$ is processed in order to produce attempted replicas of the Q encoded sub-streams. These replicas, denoted as $\{d_1(k)\}, \ldots, \{d_Q(k)\}$, are then driven to the (joint or disjoint) decoder/de-interleaver, which will attempt to recover the original uncoded sub-streams, and eventually, the original primitive bit stream. Leaving out the encoding/decoding stages, we can take an end-to-end view that relates the encoded sub-streams at the transmitter $\{b_1(k)\}, \ldots, \{b_Q(k)\}$ to their processed attempted (soft) replicas at the receiver $\{d_1(k)\}, \ldots, \{d_Q(k)\}$. Quite often, these relate in a linear fashion, that is, according to the following model:

$$\mathbf{d}(k) = \mathbf{F}\mathbf{b}(k) + \mathbf{v}(k), \qquad (3)$$

where $\mathbf{F}$ is a square $(Q \times Q)$ matrix and all vectors in (3) are of dimension $Q \times 1$. We will refer to STC super-structures that admit the end-to-end representation in (3) as *end-to-end linear*. Further, depending on the specific structure and attributes of the mixing matrix $\mathbf{F}$ and the noise impairment $\mathbf{v}(k)$, we can define some extra attributes. Before describing these attributes, we define, for convenience, a noise-prewhitened version of model (3). Denoting by $\mathbf{R_v}$ the covariance matrix of $\mathbf{v}(k)$ (i.e., $\mathbf{R_v} = E(\mathbf{v}(k)\mathbf{v}^\dagger(k))$), assumed full rank, an equivalent representation of (3) is

$$\mathbf{d}'(k) = \mathbf{F}'\mathbf{b}(k) + \mathbf{v}'(k), \qquad (4)$$

where $\mathbf{d}'(k) = \Phi_\mathbf{v}^{-1}\mathbf{d}(k)$, $\mathbf{F}' = \Phi_\mathbf{v}^{-1}\mathbf{F}$, $\mathbf{v}'(k) = \Phi_\mathbf{v}^{-1}\mathbf{v}(k)$, and $\mathbf{R_v} = E(\mathbf{v}(k)\mathbf{v}^\dagger(k)) = \Phi_\mathbf{v}\Phi_\mathbf{v}^\dagger$. Note that the new noise impairment

$\mathbf{v}'(k)$ is "spatially" white, that is, $E(\mathbf{v}'(k)\mathbf{v}'^\dagger(k)) = \mathbf{I}$. We are now ready to define a number of useful properties of end-to-end linear STC super-structures.

Decomposability

We call an end-to-end linear STC super-structure *fully decomposable*, when the matrix $\mathbf{F}'$ in (4) is diagonal (possibly after a rearrangement of its entries). In this case, the original $M \times N$ problem has reduced into Q spatially single-dimensional problems.

Partial decomposability

We call, similarly, an end-to-end linear STC super-structure *partially decomposable*, when the matrix $\mathbf{F}'$ in (4) is block-diagonal (again after a possible rearrangement of entries). In other words, instead of a coupled $Q \times Q$ problem, we are faced with a number of decoupled lower-dimension problems.

Balance

We call an end-to-end linear STC super-structure *fully balanced*, when each of the Q sub-streams in (4) experiences the same amount of interference from the other $Q - 1$ sub-streams as any other sub-stream.

Partial balance

We call an end-to-end linear STC super-structure *partially balanced*, when, the Q sub-streams can be arranged in groups of sub-streams of dimension lower than Q, such that each group experiences the same amount of interference from the other groups as any other group of sub-streams.

The features defined above, as well as some other that will be discussed later, will help us classify different STC super-structures, regarding their ability to attain their respective capacities. More precisely, they affect the way in which error correction coding can be embedded in them, so as to approach these capacities. We will now see how some of these properties and features are reflected into some particular STC super-structures.

4. $(M, 1)$ SYSTEMS

In this section, we consider some representative STC super-structures that were developed for cases of multiple-input single-output $(M, 1)$ systems. Due to the fact that MISO antenna systems provide diversity-type gains, which are, at most, logarithmic in M, (as opposed to the linear capacity increase of true MIMO systems), they are usually called "transmit diversity systems." The few techniques that will be shown are examples that fit well within the framework defined in Section 3, and as such, allow for an analytical evaluation of their theoretical constrained capacities. Before proceeding, we mention the open-loop capacity of flat-faded $(M, 1)$ system, which is given by

$$C_{M,1}^{\max} = \log_2 \left(1 + \frac{\rho}{M} \sum_{m=1}^{M} |h_m|^2 \right) \qquad (5)$$

which is obtained by substituting, in (2), $N = 1$ and $H = [h_1 \cdots h_M]$.

4.1. (2, 1) *systems: the Alamouti scheme*

An ingenious transmit diversity scheme for the $(2, 1)$ case was introduced a few years ago by Alamouti [4], and remains to date the most popular scheme for $(2, 1)$ systems. We denote by **S** the 2×2 matrix whose (i, j) element is the encoded signal going out of the jth antenna at odd $(i = 1)$ or even $(i = 2)$ time periods (the length of each time period equals the duration of one encoded symbol). In other words, one could think of the vertical dimension of **S** as representing "time" and of its horizontal dimension as representing "space." The Alamouti scheme transmits the following signal every two encoded symbol periods:

$$\mathbf{S}(k) = \begin{bmatrix} \mathbf{s}_1(k) & \mathbf{s}_2(k) \end{bmatrix} = \begin{bmatrix} b_1(k) & b_2(k) \\ b_2^*(k) & -b_1^*(k) \end{bmatrix}. \qquad (6)$$

Notice that in this case, $Q = M = 2$. Notice also that the spatial multiplexing is done according to a block scheme, the block length being equal to $L = 2$ time periods. Having assumed, as noted earlier, the channel to be flat in frequency, the $(2, 1)$ channel is characterized through $H = [h_1 \ h_2]$. The baseband signal arriving at the single receiver antenna at two consecutive time instants can be expressed as

$$\begin{aligned} \mathbf{r}(k) = \big(h_1 \big(b_1(k)\mathbf{c}_1 + b_2^*(k)\mathbf{c}_2 \big) \\ + h_2 \big(b_2(k)\mathbf{c}_1 - b_1^*(k)\mathbf{c}_2 \big) \big) + \mathbf{n}(k), \end{aligned} \qquad (7)$$

where $\mathbf{c}_1^T = [1 \mid 0]$, $\mathbf{c}_2^T = [0 \mid 1]$.[1] After sub-sampling at the receiver and complex-conjugating the second output, we obtain

$$\begin{aligned} d_1(k) = \mathbf{c}_1^T \mathbf{r}(k) = \big(h_1 b_1(k) + h_2 b_2(k) \big) + v_1(k), \\ d_2^*(k) = \mathbf{c}_2^T \mathbf{r}^*(k) = \big(-h_2^* b_1(k) \ h_1^* b_2(k) \big) + v_2(k), \end{aligned} \qquad (8)$$

where $v_m(k) = \mathbf{c}_m^T \mathbf{n}(k)$, $m = 1, 2$. Equation (8) can be equivalently written as

$$\mathbf{d}(k) = \begin{bmatrix} h_1 & h_2 \\ -h_2^* & h_1^* \end{bmatrix} \begin{bmatrix} b_1(k) \\ b_2(k) \end{bmatrix} + \mathbf{v}(k) = \mathbf{Hb}(k) + \mathbf{v}(k), \qquad (9)$$

where $\mathbf{v}^T(k) = \begin{bmatrix} v_1(k) & v_2(k) \end{bmatrix}$ and **H** is a unitary matrix. After match-filtering to **H**, we obtain

$$\begin{aligned} \mathbf{d}'(k) = \mathbf{H}^\dagger \mathbf{d}(k) \\ = \begin{bmatrix} |h_1|^2 + |h_2|^2 & 0 \\ 0 & |h_1|^2 + |h_2|^2 \end{bmatrix} \mathbf{b}(k) + \mathbf{v}'(k), \end{aligned} \qquad (10)$$

where $\mathbf{v}'(k)$ remains spatially white. Comparing to (3), it is clear that this $(2, 1)$ system is

(1) fully decomposable to two $(1, 1)$ systems;
(2) fully balanced.

The total constrained capacity of the system equals the sum of the capacities of the two SISO systems (recall that each SISO system operates at half the original information rate):

$$C_{2,1}^A = \log_2 \left(1 + \frac{\rho}{2} \big(|h_1|^2 + |h_2|^2 \big) \right). \qquad (11)$$

Note that, by contrasting (11) to (5), we see that

$$C_{2,1}^A = C_{2,1}^{\max}. \qquad (12)$$

This result is summarized in the following theorem.

Theorem 1. *The $(2, 1)$ Alamouti transmit diversity scheme has a constrained capacity equal to the $(2, 1)$ open-loop channel capacity.*[2]

Moreover, since each component of $\mathbf{v}'(k)$ is a stationary noise process, the capacity of each $(1, 1)$ system is attainable through conventional (i.e., spatially single-dimensional) state-of-the-art encoding techniques. For example, each of the two sub-streams can be encoded independently with a Turbo code, which is suitable for the classical additive white Gaussian noise channel (with stationary noise).

4.2. A (4, 1) *scheme*

The nice property of the full log-det capacity being attainable in the $(2, 1)$ case does not unfortunately hold in general for $(M, 1)$ systems with $M > 2$ (see, e.g., [10, 11]). However, some schemes have been developed recently for special cases. In the following, we describe a scheme that we recently derived for the $(4, 1)$ case (see [12]) and evaluate its capacity constrained on different receiver processing options.

The original information sequence $\tilde{b}(i)$ is first demultiplexed into four sub-streams $b_m(k)$ $(m = 1, \ldots, 4)$. The 4-dimensional transmitted signal is now organized in blocks of $L = 4$ (encoded) symbol periods, it is hence represented by a 4×4 matrix **S**, which is arranged as follows:

$$\mathbf{S} = \begin{bmatrix} b_1 & b_2 & b_3 & b_4 \\ b_2^* & -b_1^* & b_4^* & -b_3^* \\ b_3 & -b_4 & -b_1 & b_2 \\ b_4^* & b_3^* & -b_2^* & b_1^* \end{bmatrix}, \qquad (13)$$

where we have dropped the time index k for convenience. The channel matrix is again assumed flat-faded, it can be hence represented by $H = [h_1 \ h_2 \ h_2 \ h_4]$. The received signal will then be given by

$$\mathbf{r} = \mathbf{S} \begin{bmatrix} h_1 \\ h_2 \\ h_3 \\ h_4 \end{bmatrix} + \mathbf{n} = \mathbf{Sh} + \mathbf{n}, \qquad (14)$$

---

[1]By suitably redefining $\mathbf{c}_1$ and $\mathbf{c}_2$, the scheme can be modified for use with direct-sequence CDMA systems, where it is referred to as space-time spreading (STS) [9].

[2]This result has been previously reported in [6, 7].

where $r = [x(1) \quad x(2) \quad x(3) \quad x(4)]^T$ contains 4-symbol snapshots at the received signal. By complex-conjugating the second and the fourth entry of $\mathbf{r}$ in (14), we obtain

$$\mathbf{r}' = \begin{bmatrix} \mathbf{r}(1) \\ \mathbf{r}^*(2) \\ \mathbf{r}(3) \\ \mathbf{r}^*(4) \end{bmatrix} = \begin{bmatrix} h_1 & h_2 & h_3 & h_4 \\ -h_2^* & h_1^* & -h_4^* & h_3^* \\ -h_3 & h_4 & h_1 & -h_2 \\ -h_4^* & -h_3^* & h_2^* & h_1^* \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \\ b_3 \\ b_4 \end{bmatrix} + \mathbf{n}', \quad (15)$$

where $\mathbf{n}'$ is similarly obtained from $\mathbf{n}$ by complex-conjugating its second and fourth entry. The received signal is hence written as

$$\mathbf{r}' = \mathbf{Hb} + \mathbf{n}', \quad (16)$$

where now $\mathbf{H}$ is defined as

$$\mathbf{H} = \begin{bmatrix} h_1 & h_2 & h_3 & h_4 \\ -h_2^* & h_1^* & -h_4^* & h_3^* \\ -h_3 & h_4 & h_1 & -h_2 \\ -h_4^* & -h_3^* & h_2^* & h_1^* \end{bmatrix}. \quad (17)$$

We now perform matched filtering with respect to $\mathbf{H}$

$$\mathbf{r}_{\mathrm{mf}} = \mathbf{H}^\dagger \mathbf{r}' = \begin{bmatrix} \gamma & 0 & \alpha & 0 \\ 0 & \gamma & 0 & -\alpha \\ -\alpha & 0 & \gamma & 0 \\ 0 & \alpha & 0 & \gamma \end{bmatrix} \mathbf{b} + \mathbf{H}^\dagger \mathbf{n}' = \Delta_4 \mathbf{b} + \mathbf{n}_{\mathrm{mf}}, \quad (18)$$

where

$$\gamma = \mathbf{h}^\dagger \mathbf{h} = \sum_{m=1}^{4} |h_m|^2, \qquad \alpha = 2j\,\mathrm{Im}(h_1^* h_3 + h_4^* h_2). \quad (19)$$

The parameter α expresses some residual interference inherent in this $(4, 1)$ technique, and is in general nonzero. Note both the particular sparse structure of the matrix Δ_4, as well as the fact that γ is real and α is imaginary. These result in Δ_4 being in general full rank ($\det(\Delta_4) = (\gamma^2 + \alpha^2)^2$). Comparing (18) to (3), we observe that this $(4, 1)$ scheme is:

(1) partially decomposable to two uncoupled $(2, 2)$ systems;
(2) fully balanced.

Namely, by grouping the entries of $\mathbf{r}_{\mathrm{mf}}$ in two pairs, we obtain

$$\begin{bmatrix} r_{\mathrm{mf},1} \\ r_{\mathrm{mf},3} \end{bmatrix} = \Delta_2 \begin{bmatrix} b_1 \\ b_3 \end{bmatrix} + \begin{bmatrix} n_{\mathrm{mf},1} \\ n_{\mathrm{mf},3} \end{bmatrix},$$

$$\begin{bmatrix} r_{\mathrm{mf},4} \\ r_{\mathrm{mf},2} \end{bmatrix} = \Delta_2 \begin{bmatrix} b_4 \\ b_2 \end{bmatrix} + \begin{bmatrix} n_{\mathrm{mf},4} \\ n_{\mathrm{mf},2} \end{bmatrix}, \quad (20)$$

where

$$\Delta_2 = \begin{bmatrix} \gamma & \alpha \\ -\alpha & \gamma \end{bmatrix} \quad (21)$$

(we note in passing that $\Delta_2^\dagger = \Delta_2$). The two $(2, 2)$ signal models above share the same 2×2 channel matrix Δ_2 and have identically distributed, but statistically independent, 2×1 additive noise vectors. In order to facilitate the capacity evaluation of this scheme, we present at this point the noise-prewhitened version of (20):

$$\begin{bmatrix} r_1' \\ r_3' \end{bmatrix} = \Lambda \begin{bmatrix} b_1 \\ b_3 \end{bmatrix} + \begin{bmatrix} n_1' \\ n_3' \end{bmatrix}, \quad (22)$$

where

$$\Lambda = \begin{bmatrix} \lambda & \kappa \\ -\kappa & \lambda \end{bmatrix} \quad (23)$$

with $\Delta_2 = \Lambda^\dagger \Lambda$ and where n_1' and n_3' are i.i.d. and mutually independent Gaussian variables of variance σ_n^2 each. Again, an identical signal model to (22) holds for the pair $\{b_4, b_2\}$. Similar to γ and α in (21), λ and κ in (23) are real and imaginary, respectively. Similarly, the nonzero value of κ represents mutual interference between the two sub-streams (if $\alpha = 0$, then $\lambda = \sqrt{\gamma}$ and $\kappa = 0$).

Maximum allowable Shannon capacity

We first compute the Shannon capacity constrained only upon transmitter processing. By considering the two 2×2 models that describe the post-matched-filtering signals according to (22), we deduce that the maximum achievable capacity of a $(4, 1)$ system within the space-time spreading scheme (13) is given by

$$C_{4,1}^{\mathrm{constr,max}} = \frac{1}{2} \log_2 \det\left(\mathbf{I}_2 + \frac{P_T}{4\sigma_n^2} \Lambda\Lambda^\dagger \right), \quad (24)$$

where $P_T = 4\sigma_b^2$ is the total average transmitted power from the antenna array (σ_b^2 is the variance of each b_i). Since $\Lambda\Lambda^\dagger = \Delta_2$, this gives

$$C_{4,1}^{\mathrm{constr,max}} = \frac{1}{2} \log_2 \det\left(\mathbf{I}_2 + \frac{\rho}{4} \Delta_2 \right), \quad (25)$$

where $\rho = P_T/\sigma_n^2$ (all bandwidth-related normalizations have been taken into account, so that (25) represents the total capacity of the system). If the interference caused by the quantity α vanished, the expression in (25) would reduce to

$$C_{4,1}^{\mathrm{opt}} = \log_2\left(1 + \frac{\rho\gamma}{4} \right) \quad (26)$$

which is the open-loop capacity of the $(4, 1)$ flat-faded system. However, for $\alpha \neq 0$, $C_{4,1}^{\mathrm{constr,max}}$ falls short of $C_{4,1}^{\mathrm{opt}}$.

As mentioned in Section 3, the constrained capacities that correspond to each scheme depend not only on the constraints imposed at the transmitter, but also on those caused by receiver processing. In the following, we will describe a number of options for receiver processing, which will each correspond to a different constrained capacity.

4.2.1 Linear receiver processing

We observe from (20) that, in order to demodulate the transmitted sub-streams in a joint fashion, 2 input/2 output multiuser detection (MUD) is required. In this section, we present candidate receivers that perform linear MUD on each pair of matched filter outputs.

Zero-forcing processing

A straightforward way of mitigating the interference in the desired signal $\mathbf{b}$ due to α in (18), is to use a decorrelating (zero forcing—ZF) receiver. Mathematically, the ZF receiver operates on the matched-filter outputs as follows:

$$\mathbf{r}_{\mathrm{ZF}} = \Delta_4^{-1}\mathbf{r}_{\mathrm{mf}} = \mathbf{b} + \Delta_4^{-1}\mathbf{n}_{\mathrm{mf}}. \tag{27}$$

Due to the decoupling expressed in (20), the ZF operation decouples too, as follows

$$\begin{bmatrix} r_{\mathrm{ZF},1} \\ r_{\mathrm{ZF},3} \end{bmatrix} = \Delta_2^{-1}\begin{bmatrix} r_{\mathrm{mf},1} \\ r_{\mathrm{mf},3} \end{bmatrix} = \begin{bmatrix} b_1 \\ b_3 \end{bmatrix} + \Delta_2^{-1}\begin{bmatrix} n_{\mathrm{mf},1} \\ n_{\mathrm{mf},3} \end{bmatrix},$$
$$\begin{bmatrix} r_{\mathrm{ZF},4} \\ r_{\mathrm{ZF},2} \end{bmatrix} = \Delta_2^{-1}\begin{bmatrix} r_{\mathrm{mf},4} \\ r_{\mathrm{mf},2} \end{bmatrix} = \begin{bmatrix} b_4 \\ b_2 \end{bmatrix} + \Delta_2^{-1}\begin{bmatrix} n_{\mathrm{mf},4} \\ n_{\mathrm{mf},2} \end{bmatrix}. \tag{28}$$

Note that (28) is equivalent to (27). The ZF receiver detects the four sub-streams by further processing the zero-forcing outputs, that is, the entries of the vector $\mathbf{r}_{\mathrm{ZF}}$ given in (27). Each of the four zero-forcing outputs can be seen as the output of the following AWGN channel:

$$r_{\mathrm{ZF},i} = b_i + n_{\mathrm{ZF},i}, \quad i = 1,\ldots,4, \tag{29}$$

where $n_{\mathrm{ZF},i}$ is an i.i.d. Gaussian noise independent of b_i, of variance that can be found to equal $\gamma\sigma_n^2/(\gamma^2 + \alpha^2)$. At this point, the system has been reduced to a fully decomposed, fully balanced system. Hence, its capacity is given by

$$C_{4,1}^{\mathrm{ZF}} = \log_2\left(1 + \frac{\rho}{4}\left(\frac{\gamma^2 + \alpha^2}{\gamma}\right)\right) \quad [\mathrm{bps/Hz}], \tag{30}$$

where we recall that $\rho = P_T/\sigma_n^2 = 4\sigma_b^2/\sigma_n^2$. Notice that the four sub-streams have equal capacities, and that the total capacity of the system equals four times that of any given sub-stream.

MMSE processing

A better compromise between signal recovery and noise amplification (and hence better performance) can be achieved with minimum mean squared error (MMSE) processing. This is achieved by the 4×4 setting $W_{\mathrm{MS},4}$ which minimizes the MMSE criterion:

$$\min_{W_{\mathrm{MS},4}} \left\| W_{\mathrm{MS},4}^\dagger \mathbf{r}_{\mathrm{mf}} - \mathbf{b} \right\|^2. \tag{31}$$

The minimization of (31) yields the Wiener solution

$$W_{\mathrm{MS},4}^\dagger = \Delta_4^\dagger\left(\Delta_4\Delta_4^\dagger + \frac{\sigma_n^2}{\sigma_b^2}\mathbf{H}^\dagger\mathbf{H}\right)^{-1}. \tag{32}$$

Hence, the post-MMSE-processed signal delivered to the detector is given by

$$\mathbf{r}_{\mathrm{MS}} = \Delta_4^\dagger\left(\Delta_4\Delta_4^\dagger + \frac{\sigma_n^2}{\sigma_b^2}\mathbf{H}^\dagger\mathbf{H}\right)^{-1}\mathbf{r}_{\mathrm{mf}}$$
$$= \Delta_4^\dagger\left(\Delta_4\Delta_4^\dagger + \frac{\sigma_n^2}{\sigma_b^2}\Delta_4\right)^{-1}\mathbf{r}_{\mathrm{mf}}. \tag{33}$$

Similar to the ZF case, the MMSE solution in (33) is, similar to (28), decomposable as follows:[3]

$$\begin{bmatrix} r_{\mathrm{MS},1} \\ r_{\mathrm{MS},3} \end{bmatrix} = \Delta_2^\dagger\left(\Delta_2\Delta_2^\dagger + \frac{\sigma_n^2}{\sigma_b^2}\Delta_2\right)^{-1}\begin{bmatrix} r_{\mathrm{mf},1} \\ r_{\mathrm{mf},3} \end{bmatrix} = W_{\mathrm{MS},2}^\dagger\begin{bmatrix} r_{\mathrm{mf},1} \\ r_{\mathrm{mf},3} \end{bmatrix},$$
$$\begin{bmatrix} r_{\mathrm{MS},4} \\ r_{\mathrm{MS},2} \end{bmatrix} = \Delta_2^\dagger\left(\Delta_2\Delta_2^\dagger + \frac{\sigma_n^2}{\sigma_b^2}\Delta_2\right)^{-1}\begin{bmatrix} r_{\mathrm{mf},4} \\ r_{\mathrm{mf},2} \end{bmatrix} = W_{\mathrm{MS},2}^\dagger\begin{bmatrix} \mathbf{r}_{\mathrm{mf},4} \\ \mathbf{r}_{\mathrm{mf},2} \end{bmatrix}. \tag{34}$$

In this case too, an equation similar to (29) can be written, wherein each sub-stream is detected at the output of a $(1, 1)$ system and AWGN noise. The additive noise will contain now contributions from one other sub-stream, it has however the same variance for all four sub-streams. So again, the system has been fully decomposed to four $(1, 1)$ systems in a fully balanced way. It is then straightforward to compute the capacity of the MMSE receiver, which is given by

$$C_{4,1}^{\mathrm{MMSE}} = \log_2\left(1 + \frac{W_1^\dagger\Omega W_1}{W_1^\dagger\Phi W_1 + 4/\rho\left(W_1^\dagger\Delta_2 W_1\right)}\right), \tag{35}$$

where

$$\Omega = \begin{bmatrix} \gamma \\ -\alpha \end{bmatrix}\begin{bmatrix} \gamma \\ -\alpha \end{bmatrix}^\dagger, \quad \Phi = \begin{bmatrix} \alpha \\ \gamma \end{bmatrix}\begin{bmatrix} \alpha \\ \gamma \end{bmatrix}^\dagger, \quad W_1^\dagger = \begin{bmatrix} 1 & 0 \end{bmatrix}W_{\mathrm{MS},2}^\dagger, \tag{36}$$

where $W_{\mathrm{MS},2}^\dagger$ is given in (34).

4.2.2 Maximum likelihood MUD

We now focus on the prewhitened signal model (22), which we repeat here for convenience:

$$\begin{bmatrix} r_1' \\ r_3' \end{bmatrix} = \Lambda\begin{bmatrix} b_1 \\ b_3 \end{bmatrix} + \begin{bmatrix} n_1' \\ n_3' \end{bmatrix}. \tag{37}$$

Keeping in mind that the noise vector $\begin{bmatrix} n_1' & n_3' \end{bmatrix}^T$ is jointly Gaussian with covariance matrix $\sigma_n^2\mathbf{I}_2$, the maximum likelihood (ML) multiuser detector for (37) solves the following optimization problem:

---

[3]As expected, as $\rho \to \infty$, the solution $W_{\mathrm{MS},2}^\dagger$ in (34) converges to the ZF solution $W_{\mathrm{ZF},2}^\dagger = \Delta_2^{-1}$.

$$\min_{\{b_1, b_3\} \in \mathscr{A} \times \mathscr{A}} \left\| \begin{bmatrix} r_1' \\ r_3' \end{bmatrix} - \Lambda \begin{bmatrix} b_1 \\ b_3 \end{bmatrix} \right\|^2, \tag{38}$$

where $\mathscr{A}$ is the alphabet shared by all the encoded sub-streams.

Equation (38) is a typical maximum likelihood MUD problem (see [13]). Typically, in order to avoid an exhaustive multi-dimensional search, the encoding imparts a special structure (such as with convolutional codes). Then the use of dynamic programming techniques such as the Viterbi Algorithm (VA) provides an important saving in complexity.

We are now ready to assess the capacity of the proposed $(4, 1)$ super-structure, constrained on ML reception. Consider a pair of transmitted sequences $\{\tilde{b}_1\}$, $\{\tilde{b}_3\}$, to be encoded in a spatially balanced way (either independently or jointly). Then, because of the symmetrical structure of the channel matrix Λ in (37), the communication system is perfectly balanced (it is understood that the encoding of each sequence at the transmitter is done without knowledge of the channel instantiations). A spatially two-dimensional version of Shannon's classical random coding procedure then applies. We start with a primitive (maxentropic) independent bit stream, and demultiplex it into its even and odd sub-streams, b_1 and b_3, respectively. The encoded sequence is assigned half of its bits (b_1) from the first sub-stream (first dimension), and the other half (b_3) from the second sub-stream (second dimension). Then, due to the perfect balance between the two dimensions, the system's capacity is achieved when each of the two sub-streams achieves its own (half of the full) capacity. This requires, however, joint optimal (minimum distance) detection of the two sub-streams, as per (38).

In conclusion, the Shannon capacity which is achieved through ML detection (in the limit of infinitely long random codes) is given by (25), which we repeat here for convenience

$$C_{4,1}^{\mathrm{ML}} = \frac{1}{2} \log_2 \det \left(\mathbf{I}_2 + \frac{\rho}{4} \Delta_2 \right). \tag{39}$$

As will be shown later, this capacity is very close to the full $(4, 1)$ capacity $C_{4,1}^{\max}$. Notice that in the ML case, even though it is fully balanced, the system has been only partially decomposed in two 2×2 systems.

4.3. Nonfull rate $(4, 1)$ codes

We now describe some easily derived constrained capacities of some other, less optimal, but quite simple, $(M, 1)$ schemes.

(1) STS(3,1)—3/4 rate: in [9] it was shown that a $(3, 1)$ scheme can be designed, which achieves the full $(3, 1)$ capacity, but at the price of a 25% loss of rate. This scheme uses block multiplexing, in a fashion similar to the above schemes of Section 4. It multiplexes $Q = 3$ sub-streams on 3 antennas, over $L = 4$ symbol periods. It results, however, in a fully-decomposable, fully-balanced, 3×3 system with stationary noise. Its constrained capacity is given by

$$C_{31-3/4} = \frac{3}{4} \log_2 \left(1 + \frac{\rho(|h_1|^2 + |h_2|^2 + |h_3|^2)}{3} \right). \tag{40}$$

(2) STS(4, 1)—real: it was also mentioned in [9] and elsewhere, that a fully decomposable and fully balanced extension of the Alamouti $(2, 1)$ scheme for real inputs can be used for a $(4, 1)$ system. In the case of complex inputs, it is possible to use the same scheme if we sacrifice 50% of the rate, that is by signaling half of the time on each complex dimension. The capacity of this scheme equals half of the $(4, 1)$ open-loop capacity:

$$C_{41-\mathrm{real}} = \frac{1}{2} C_{4,1}^{\max}. \tag{41}$$

4.4. $(M, 1)$ hopping

A very simple alternative that can be used for any integer M is based on the idea of cycling a single encoded stream over the four transmit antennas. In this case, the data stream is first encoded as a single stream $\{b(k)\}$. The encoded sequence $\{b(k)\}$ is then demultiplexed into M sub-sequences, $\{b_1(k)\}, \dots, \{b_M(k)\}$. The mth subsequence is transmitted from the mth antenna ($m = 1, \dots, M$). In other words, the M antennas take turns in transmitting (at full power) the M sub-streams of the single encoded data sequence. This scheme is fully balanced and fully decomposable, however its noise impairment is not stationary. Its capacity is easily found to be given by the average capacity of the M full-power $(1,1)$ sub-channels, that is,

$$C_{M,1}^{\mathrm{hop}} = \frac{1}{M} \sum_{m=1}^{M} \log_2 \left(1 + \rho |h_m|^2 \right). \tag{42}$$

It is important to emphasize that a peculiarity of this simple approach is that the encoded sequence is effectively transmitted through a channel whose SNR is periodic. Conventional encoding techniques do not perform in general satisfactorily with such periodic channels. Special codes that can cope with such channels are required in order to be able to approach the capacity in (42). These codes are a current research topic [14].

Discussion

Other approaches for the $(M, 1)$ case have appeared in the recent literature. An exhaustive listing of all of them would be however beyond the scope of this paper. We should note further that the benefit of open-loop $(M, 1)$ systems becomes increasingly limited as M grows. Keeping the total transmit power from all the antennas constant, assuming that $E|h_m|^2 = 1$ and letting M to grow towards infinity, the $(M, 1)$ open-loop capacity in (5) tends to the following asymptote:

$$C_{\infty,1} = \log_2(1 + \rho). \tag{43}$$

It is clear from (43) that we cannot keep increasing the capacity of an open-loop $(M, 1)$ open-loop system by simply increasing the number of transmitter antennas. The use of more antennas at the receiver becomes necessary when higher capacities are sought.

5. (M, N) SYSTEMS

In this section, we analyze some STC architectural super-structures for the case of $N > 1$ receiver antennas.

5.1.   *Combined transmit/receive diversity systems*

Given a certain $(M, 1)$ system, one straightforward way to design an (M, N) system is to simply:

- transmit as in the $(M, 1)$ system,
- receive on each antenna as in the $(M, 1)$ system,
- combine optimally the N receiver antenna outputs.

The capacity quantification of these transmit/receive diversity systems is straightforward. The $M \times N$ (assumed flat) channel is represented through the $N \times M$ channel matrix

$$\mathbf{H} = \begin{bmatrix} h_{11} & \cdots & h_{1M} \\ \vdots & \ddots & \vdots \\ h_{N1} & \cdots & h_{NM} \end{bmatrix} = \begin{bmatrix} \mathbf{h}_1 & \cdots & \mathbf{h}_N \end{bmatrix}. \qquad (44)$$

We first compute an upper bound for the capacity of such an (M, N) transmit/receive diversity system. With optimal ratio combining, and assuming that each $(M, 1)$ system takes no interference hit, the input/output relationship takes the form

$$\mathbf{d}(k) = \left(\sum_{m=1}^{M} \sum_{n=1}^{N} |h_{nm}|^2 \right) \mathbf{b}(k) + \mathbf{n}(k) \qquad (45)$$

corresponding to the capacity

$$C_{M,N}^{\text{trd,max}} = \log_2 \left(1 + \frac{\rho}{M} \sum_{m=1}^{M} \sum_{n=1}^{N} |h_{nm}|^2 \right). \qquad (46)$$

It is clear that, when the attainable capacity of the corresponding $(M, 1)$ schemes is away from the $(M, 1)$ log-det capacity, the upper bound in (46) will not be attained either. Notice further that, the expression in (46) is strictly smaller than the (M, N) log-det capacity in (2) for $N > 1$.

Examples

To give some examples, the capacity of a $(2, N)$ system that uses the Alamouti $(2, 1)$ super-structure is

$$C_{2,N}^{A} = \log_2 \left(1 + \frac{\rho}{2} \sum_{n=1}^{N} \left(|h_{n,1}|^2 + |h_{n,2}|^2 \right) \right), \qquad (47)$$

that is, as expected, the upper bound in (46) is attained by the Alamouti scheme in the $(2, N)$ case.

It is also straightforward to compute the maximum attainable capacity of a $(4, N)$ system that uses the $(4, 1)$ scheme of Section 4.2, which is given by

$$C_{4,N}^{\text{ML}} = \frac{1}{2} \log_2 \det \left(\mathbf{I}_2 + \frac{\rho}{4} \Lambda_{2n} \Lambda_{2n}^{\dagger} \right), \qquad (48)$$

where

$$\Lambda_{2n} = \begin{bmatrix} \Lambda_1 \\ \vdots \\ \Lambda_n \end{bmatrix} \qquad (49)$$

and Λ_n is defined, similar to (23) for the nth receiver antenna.

5.2.   *V-BLAST*

A quite simple, from the transmitter's point of view, STC super-structure was proposed in [8], and is widely referred-to as "V-BLAST." In this architecture, $\{(bi)\}$ is first demultiplexed into M sub-streams, which are then encoded independently and mapped each on a different antenna:

$$s_m(k) = b_m(k). \qquad (50)$$

In other words, the original bit stream is converted into a vertical vector of encoded sub-streams (whence the term "vertical" BLAST) which are then streamed to the antennas through a 1-1 mapping. In [8], it was proposed to process the received signal with the use of a successive interference canceller. After determining the order into which the M sub-streams will be detected, the V-BLAST receiver operates according to the following generic 3-stage scheme, which is followed in a successive fashion for each sub-stream:

(1) project away from the remaining interfering sub-streams;
(2) detect (after de-coding, de-interleaving, and slicing) the sub-stream;
(3) cancel the effect of the detected sub-stream from subsequent sub-streams.

Mathematically, these operations can be described as follows for the k_mth sub-stream:

$$\begin{aligned} z_{k_m}(k) &= W_{k_m}^{\dagger} \mathbf{x}^m(k), \\ \hat{z}_{k_m}(k) &= \text{dec}(z_{k_m}(k)), \\ \mathbf{x}^{m+1}(k) &= \mathbf{x}^m(k) - \text{enc}(\hat{z}_{k_m}(k)) \mathbf{h}_{k_m}, \end{aligned} \qquad (51)$$

where $\mathbf{x}^1(k) = \mathbf{x}(k)$, $\{k_1, \ldots, k_M\}$ is a reordering of the set $\{1, \ldots, M\}$ that determines the order in which the sub-streams will be detected, $\text{dec}(\cdot)$ represents the decoding plus detection operation, and $\text{enc}(\cdot)$ represents the encoding operation. Finally, W_{k_m} represents the $N \times 1$ vector that operates on $\mathbf{x}^m(k)$ in order to project away from sub-streams $\{k_{m+1}, \ldots, k_M\}$. The operations in (51) are performed successively for $m = 1, \ldots, M$, after the ordering $\{k_1, \ldots, k_M\}$ has been determined.

We now discern between the following two cases for this linear operation, since they affect significantly the constrained capacity of the system.

Zero-forcing projection

In this case, at the mth stage, $W_{k_m}^{\dagger}$ nulls perfectly the interference from all the remaining (undetected) sub-streams. These

are the sub-streams with indices $\{k_{m+1}, \ldots, k_M\}$. This nulling is represented mathematically as

$$W_{\text{ZF},k_m}^\dagger \mathbf{H} = \begin{bmatrix} 0 & \cdots & 0 & 1 & 0 & \cdots & 0 \end{bmatrix} = \delta_{k_m}, \qquad (52)$$

where the unique nonzero element of the $1 \times M$ vector δ_{k_m} is in its k_mth position. As a result, the end-to-end model for the k_mth output is

$$d_{k_m}(k) = b_{k_m}(k) + W_{\text{ZF},k_m}^\dagger \mathbf{n}(k), \quad m = 1, \ldots, M, \qquad (53)$$

where $\mathbf{n}(k) = \begin{bmatrix} n_1(k) & \cdots & n_N(k) \end{bmatrix}^T$ is the receiver noise. Defining

$$\begin{aligned} \mathbf{d}(k) &= \begin{bmatrix} d_{k_1}(k) & \cdots & d_{k_M}(k) \end{bmatrix}^T, \\ \mathbf{b}(k) &= \begin{bmatrix} b_{k_1}(k) & \cdots & b_{k_M}(k) \end{bmatrix}^T, \end{aligned} \qquad (54)$$

equation (53) can be written in matrix form as

$$\mathbf{d}(k) = \mathbf{b}(k) + \mathbf{W}_{\text{ZF}}^\dagger \mathbf{n}(k), \qquad (55)$$

where $\mathbf{W}_{\text{ZF}} = \begin{bmatrix} W_{\text{ZF},k_1} & \cdots & W_{\text{ZF},k_M} \end{bmatrix}$. From (55), it is obvious that the ZF version of the V-BLAST super-structure is a fully decomposable, however *not fully balanced* system, due to the generally different square norms of the different columns of $\mathbf{W}_{\text{ZF}}$.

Regarding the capacity of the end-to-end system, it is important to emphasize that we have assumed that each sub-stream is independently encoded, and that the transmitter has no way of knowing which is the highest rate for each antenna. As a result, it can only transmit from all antennas the same rate. Hence, the capacity will equal M times the smallest of the M decomposed channel capacities:

$$C_{\text{MN}}^{\text{VB-ZF}} = M \times \min_{m \in \{1, \ldots, M\}} \left\{ \log_2 \left(1 + \rho_{\text{ZF},k_m} \right) \right\}, \qquad (56)$$

where ρ_{k_m} is the output SNR of the k_mth sub-stream:

$$\rho_{\text{ZF},k_m} = \frac{\rho}{M \| W_{\text{ZF},k_M} \|^2}. \qquad (57)$$

It should finally be noted that the capacity in (56) can be optimized by choosing an optimal ordering for the set $\{k_1, \ldots, k_M\}$ (see [8]).

MMSE projection

In this case, at the mth stage, an optimal compromise between linear interference mitigation of the undetected sub-streams and noise amplification is sought. This is achieved through the following MMSE criterion:

$$\min_{W_{k_m}} E \| d_{k_m} - W_{k_m}^\dagger \mathbf{H}_{k_m} \|^2, \qquad (58)$$

where $\mathbf{H}_{k_m}$ is derived from $\mathbf{H}$ by deleting its columns corresponding to indices $\{k_1, \ldots, k_{m-1}\}$. This gives for W_{k_m}:

$$W_{\text{MMSE},k_m}^\dagger = \left(\mathbf{H}_{k_m} \mathbf{H}_{k_m}^\dagger + \frac{M}{\rho} \mathbf{I}_N \right)^{-1} \mathbf{h}_{k_m}, \qquad (59)$$

where $\mathbf{h}_{k_m}$ is the k_mth column of $\mathbf{H}$. This end-to-end system has now been fully decomposed into M $(1, 1)$ systems, however, it is not generally balanced. Its capacity is hence computed again through the minimum of the M 1×1 capacities (assuming Gaussian signaling for each sub-stream), and is given by a formula similar to (56):

$$C_{\text{MN}}^{\text{VB-MMSE}} = M \times \min_{m \in \{1, \ldots, M\}} \left\{ \log_2 \left(1 + \rho_{\text{MMSE},k_m} \right) \right\}, \qquad (60)$$

where now

$$\rho_{\text{MMSE},k_m} = \frac{\| W_{\text{MMSE},k_m}^\dagger \mathbf{H}_{k_m} \|^2}{M \| W_{\text{MMSE},k_m} \|^2 / \rho + \sum_{l \neq k_m} \| W_{\text{MMSE},l} \|^2}. \qquad (61)$$

Again, the capacity in (60) can be maximized through optimal ordering.

5.3. Other (M, N) schemes

Similar to the $(M, 1)$ case, several other schemes have been proposed in the literature for the general (M, N) case. For example, it was suggested in [15] to use a block space-time multiplexing whose mixing coefficients are derived numerically according to a maximum average capacity criterion. Another approach in [16] uses Turbo codes in the following way: the original sub-stream is first demultiplexed into M sub-streams, which are separately encoded each with a block code. Then, the M encoded outputs are space-time interleaved in a random fashion, mapped onto constellation symbols, and sent out of the M antennas. At the receiver, the M sub-streams are separated through an iterative interference canceller, which uses MMSE for the linear (soft) part, and subtracts decisions made after (joint) de-interleaving and (separate) decoding of each interfering sub-stream in the cancellation part.

These approaches have demonstrated encouraging performance in terms of bit/frame error rate at the receiver. However, their inherent capacity penalties are still unknown, due mainly to their apparent luck of structure and other properties such as the ones discussed above. The quantification of the capacity penalties of these and other emerging STC super-structures remains an interesting open question.

6. NUMERICAL RESULTS

We will now show some representative capacity plots for the STC architectures that were mentioned above. In all cases, we will use the analytical expressions derived in the paper. We will run these expressions over an ensemble of (M, N) random Rayleigh-faded channel matrices (each entry of the matrix is chosen independently from any other entry from a complex i.i.d. Gaussian distribution of unit variance). We will then plot *outage* capacities, that is, we will pick out of the computed capacity cdf a point according to a typical outage percentage (such as 10%, which is typical in wireless communications).
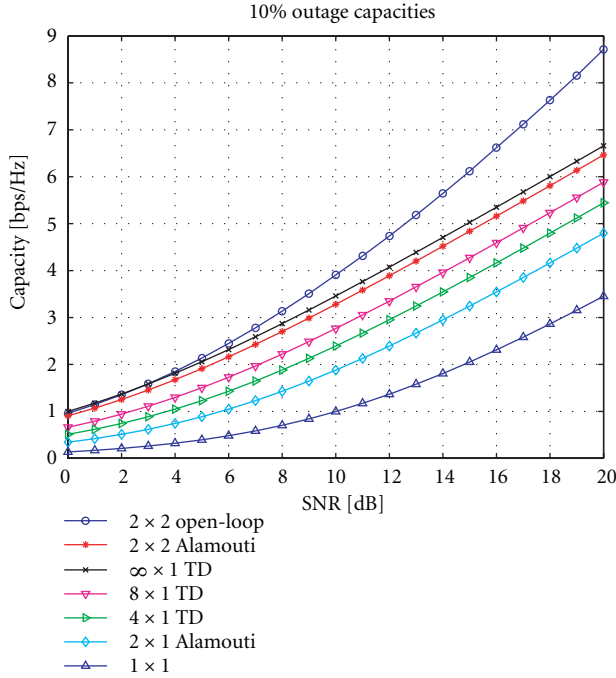
FIGURE 2: Outage capacities and bounds of $(M, 1)$ and $(M, 2)$ schemes.

In Figure 2, we show the 10% outage capacities for several $(M, 1)$ cases, as well as for the $(2, 2)$ case. In the $(2, 1)$ case, the plotted capacity corresponds to both the Alamouti scheme and to the maximum open-loop capacity, as indicated by (12). For the other $(M, 1)$ cases, we plot the capacity upper bounds corresponding to (5), and we use (43) for the asymptotic $(\infty, 1)$ case. We also use (46) with $N = 2$ for the capacity of a $(2, 2)$ combined Alamouti/receive diversity scheme, and the log-det expression (2) for the $(2, 2)$ maximum open-loop capacity. We observe that, at $\rho = 10$ dB, the $(2, 1)$ system almost doubles the capacity of the $(1, 1)$ system! However, as noted earlier, increasing the number of transmit antennas in the $(M, 1)$ case offers diminishing returns. It is also worth noting that the $(2, 2)$ combined transmit/receiver diversity scheme is capable of attaining a quite significant fraction (particularly at low SNR's) of the maximum $(2, 2)$ open-loop capacity. Finally, it is also interesting to note that a $(2, 2)$ system achieves about the same capacity as a $(\infty, 1)$ system, which conveys again the message of the high value of adding extra antennas at the receiver.

In Figure 3, we show the capacities of the $(4, 1)$ scheme of Section 4.2 when used in conjunction with the different proposed receiver architectures (ZF, MMSE, and ML). It is noticeable that the ML structure approaches closely the channel's $(4, 1)$ open-loop capacity. Moreover, we observe that at low SNR's, the linear MMSE solution is also very close to the open-loop capacity. Table 1 shows some of these results at chosen SNR points.

Figure 4 shows a comparison of some of the $(M, 1)$ systems mentioned in Section 4, including non-full rate variants of the Alamouti (STS) scheme. We observe that the non-full rate schemes fall well behind the $(2, 1)$ scheme in terms of
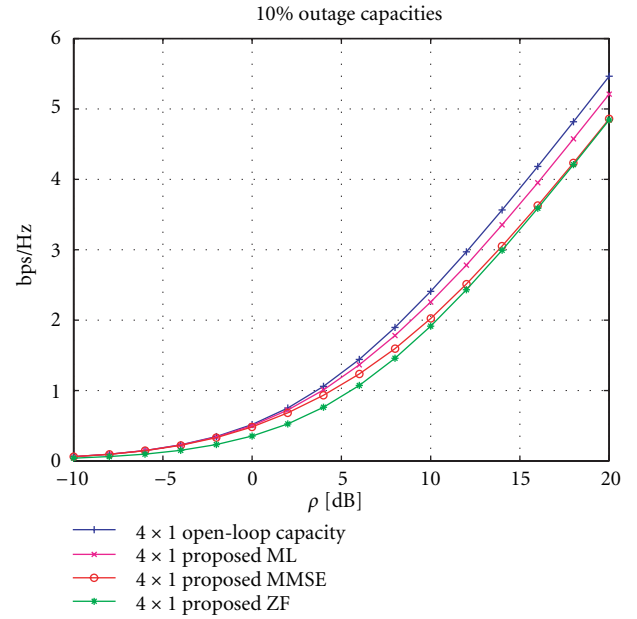

FIGURE 3: Outage capacities of the $(4, 1)$ scheme of Section 4.2 compared to the $(4, 1)$ open-loop capacity.

TABLE 1: Indicative outage capacities of the proposed $(4, 1)$ scheme versus the $(4, 1)$ open-loop capacity.

ρ [dB]	ZF	MMSE	ML	OPT
-10	0.038	0.056	0.057	0.057
0	0.344	0.469	0.480	0.491
10	1.886	1.990	2.212	2.339
20	4.805	4.825	5.155	5.379

outage capacity. Moreover, the $(2, 1)$ scheme is increasingly close to the $(4, 1)$ open-loop capacity at low SNR's. Similarly, in Figure 5, we show comparative plots of the capacity of the $(4, 1)$ hopping scheme mentiond in Section 4.2, (see (42)).

In Figure 6, we show the capacities of some combined transmit/receive diversity schemes for different $(4, N)$ cases. The circles represent the combined $(4, N)$ systems corresponding to the $(4, 1)$ scheme of Section 4.2, in conjunction with optimal receiver diversity. When read from the bottom up, these four curves correspond to $N = 1, 2, 3, 4$, respectively. Similarly, the crosses represent the corresponding open loop $(4, N)$ capacities. Notice that the proposed $(4, 1)$ scheme is very close to the open-loop capacity, however the gap gets increasingly larger as N grows from 1 to 4. In the $(4, 2)$ case however, the scheme still performs very well at low SNR's.

Finally, in Figure 7, we show a capacity cdf, at 10 dB SNR, of the ZF and MMSE V-BLAST architectures described in Section 5.2 for the $(4, 4)$ case. Notice that, at this SNR, the MMSE architecture is able of attaining about 70% of the total open-loop capacity at 10% outage. However, the ZF architecture performs poorly, and is even outperformed by a $(1, 4)$ maximal ratio combining system at outages smaller than 20%! The situation is more severe for lower SNR's such as 0 dB, as shown in Figure 8. Now the V-BLAST MMSE
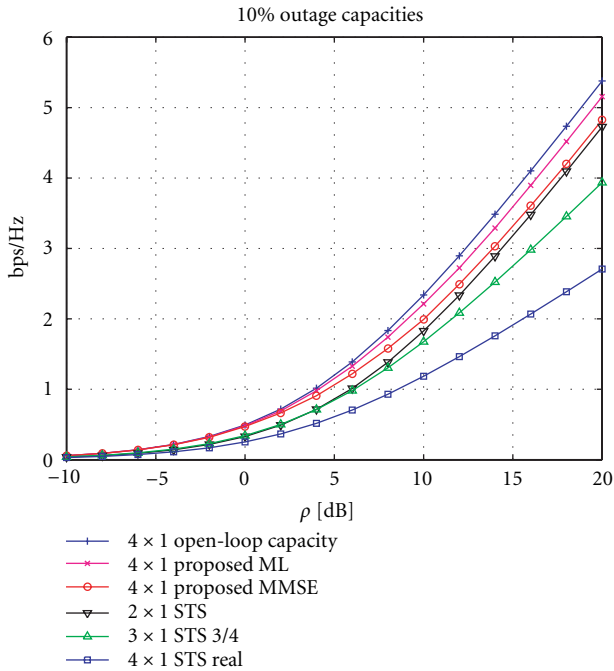
FIGURE 4: 10% outage capacities compared to other open-loop alternatives.
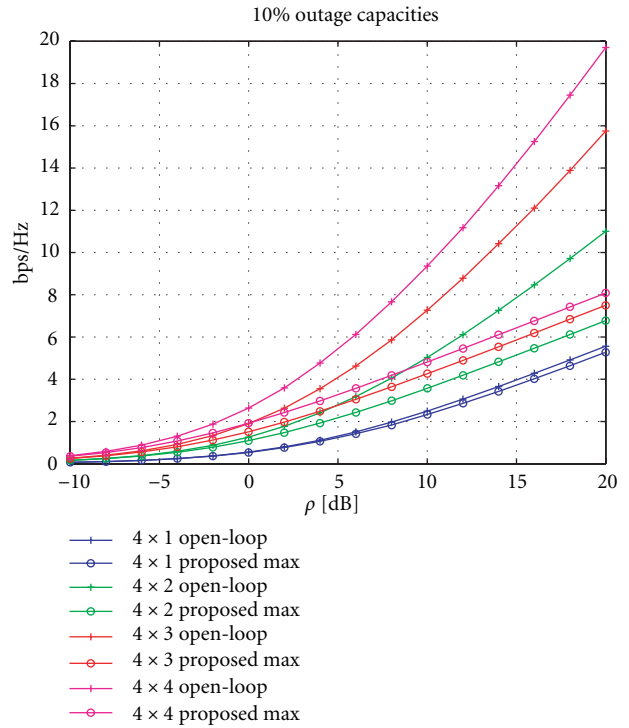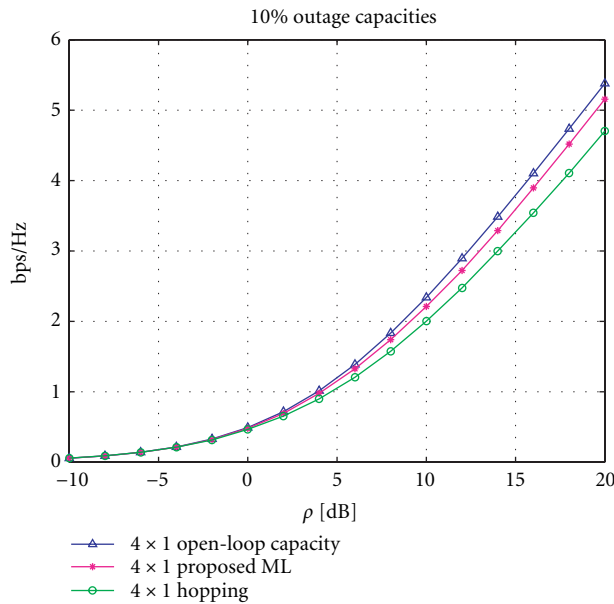

FIGURE 5: Outage capacity of a hopping scheme.

architecture attains only about 50% of the $(4, 4)$ open-loop capacity, whereas the ZF architecture is outperformed by the $(1, 4)$ system across the board.

7. CONCLUSIONS

We have presented a framework for analyzing space-time coding architectures in terms of Shannon capacity. We defined a number of attributes of such schemes that allow

FIGURE 6: Outage capacities of the $(4, 1)$ scheme described in Section 4.2, when used with up to four receiver antennas.
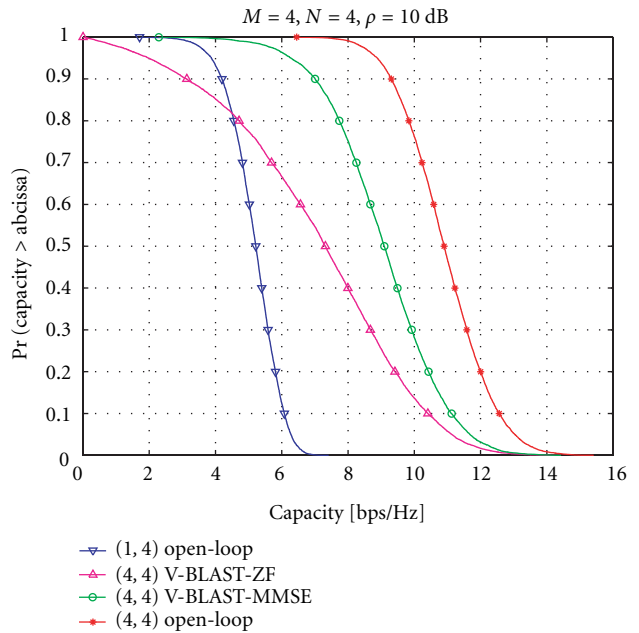

FIGURE 7: Outage capacity distribution of a V-BLAST MMSE architecture at 10 dB SNR.

for the computation of their inherent capacity penalties, which we have computed analytically for a few representative examples. Our theoretically derived expressions were also numerically validated by a number of computer simulations that compare the considered architectures in terms of outage
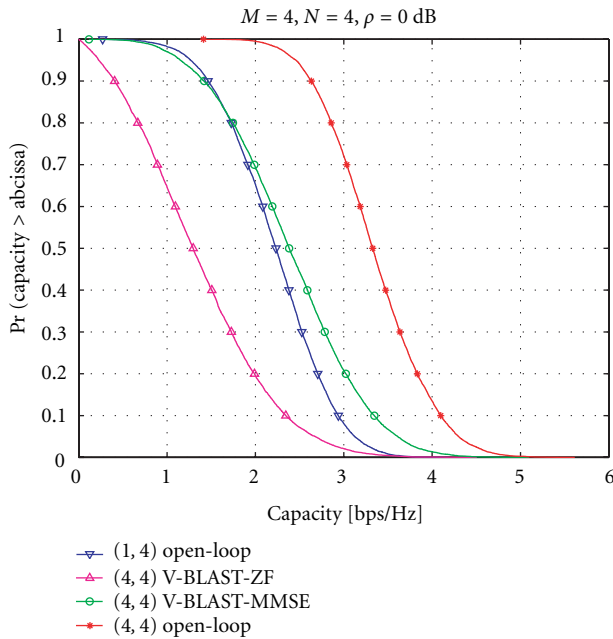
FIGURE 8: Outage capacity distribution of a V-BLAST MMSE architecture at 0 dB SNR.

capacity. We believe that these results provide some useful intuition regarding the performance trade-offs of different techniques. Future work will be targeted in analyzing other promising STC schemes, as well as in determining new architectures of higher capacity potential.

REFERENCES

[1] G. J. Foschini, "Layered space-time architecture for wireless communication in a fading environment when using multi-element antennas," *Bell Labs Technical Journal*, vol. 1, no. 2, pp. 41–59, 1996.

[2] C. Berrou, A. Glavieux, and P. Thitimajshima, "Near Shannon-limit error correction coding and decoding: Turbo codes," in *Proc. 1993 International Conference on Communications*, pp. 1064–1070, Geneva, Switzerland, May 1993.

[3] C. Shannon, "A mathematical theory of communication," *Bell System Technical Journal*, vol. 27, pp. 379–423, 623–656, 1948.

[4] S. Alamouti, "A simple transmitter diversity scheme for wireless communications," *IEEE Journal on Selected Areas in Communications*, vol. 16, no. 8, pp. 1451–1458, 1998.

[5] V. Tarokh, N. Seshadri, and A. R. Calderbank, "Space-time codes for high data rate wireless communication: performance criterion and code construction," *IEEE Transactions on Information Theory*, vol. 44, no. 2, pp. 744–765, 1998.

[6] C. Papadias, "On the spectral efficiency of space-time spreading schemes for multiple antenna CDMA systems," in *33rd Asilomar Conference on Signals, Systems, and Computers*, pp. 639–643, Pacific Grove, Calif, USA, October 1999.

[7] S. Sandhu and A. Paulraj, "Space-time block codes: a capacity perspective," *IEEE Communications Letters*, vol. 4, no. 12, pp. 384–386, 2000.

[8] G. J. Foschini, G. D. Golden, R. A. Valenzuela, and P. W. Wolniansky, "Simplified processing for wireless communication at high spectral efficiency," *IEEE Journal on Selected Areas in Communications*, vol. 17, no. 11, pp. 1841–1852, 1999.

[9] B. Hochwald, L. Marzetta, and C. Papadias, "A transmitter diversity scheme for wideband CDMA systems based on space-time spreading," *IEEE Journal on Selected Areas in Communications*, vol. 19, no. 1, pp. 48–60, 2001.

[10] A. V. Geramita and J. Seberry, *Orthogonal Designs: Quadratic Forms and Hadamard Matrices*, Marcel Dekker, New York, USA, 1979.

[11] G. Ganesan and P. Stoica, "Space-time diversity using orthogonal and amicable orthogonal designs," in *Proc. IEEE Int. Conf. Acoustics, Speech, Signal Processing*, Istanbul, Turkey, June 2000.

[12] C. Papadias and G. J. Foschini, "A space-time coding approach for systems employing four transmit antennas," in *International Conference on Acoustics, Speech, and Signal Processing*, Salt Lake City, Utah, USA, May 2001.

[13] S. Verdú, *Multi-User Detection*, Cambridge University Press, Cambridge, UK, 1999.

[14] R. D. Wesel, X. Liu, and W. Shi, "Trellis codes for periodic erasures," *IEEE Trans. Communications*, vol. 48, no. 6, pp. 938–974, 2000.

[15] B. Hassibi and B. Hochwald, "High-rate linear space-time codes," in *Proc. IEEE Int. Conf. Acoustics, Speech, Signal Processing*, Salt Lake City, Utah, USA, May 2001.

[16] M. Sellathurai and S. Haykin, "Joint beamformer estimation and co-antenna interference cancelation for turbo-BLAST," in *Proc. IEEE Int. Conf. Acoustics, Speech, Signal Processing*, Salt Lake City, Utah, USA, May 2001.

Constantinos B. Papadias was born in Athens, Greece, in 1969. He received the diploma of electrical engineering from the National Technical University of Athens (NTUA) in 1991 and the Ph.D. degree in signal processing (highest honors) from the Ecole Nationale Supérieure des Télécommunications (ENST), Paris, France, in 1995. From 1992 to 1995, he was a Teaching and Research Assistant at the Mobile Communications Department, Eurécom, France. In 1995, he joined the Information Systems Laboratory, Stanford University, Stanford, Calif, USA, as a PostDoctoral Researcher, working in the Smart Antennas Research Group. In November 1997, he joined the Wireless Research Laboratory of Bell Labs, Lucent Technologies, Holmdel, NJ, USA, as a Member of Technical Staff. He now heads the Global Wireless Strategy Research group in the same lab. His current research interests lie in the areas of multiple antenna systems (e.g., MIMO transceiver design, and space-time coding), interference mitigation techniques, reconfigurable wireless networks, as well as financial evaluation of wireless technologies. He has authored several papers and patents on these topics. Dr. Papadias is a member of the Technical Chamber of Greece.

Gerard J. Foschini BSEE-NJIT, MEE-NYU, Ph.D. Mathematics-Stevens. Mr. Gerard J. Foschini has been at Bell Laboratories for nearly 40 years. He holds the position of Distinguished Member of Staff. He has conducted data communications research on many kinds of systems, most recently wireless communications and optical communications systems. Gerard has done extensive research on point to point systems as well as on networks. He won the 2001 Bell Labs, Distinguished Inventor Award. He is an IEEE Fellow and he has taught at Princeton and Rutgers.